Energy oriented Center of Excellence for computing applications (EoCoE):

Performance evaluation and optimisation methodology

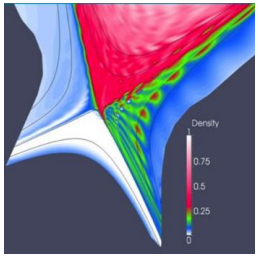
Paul Gibbon
Forschungszentrum Jülich GmbH

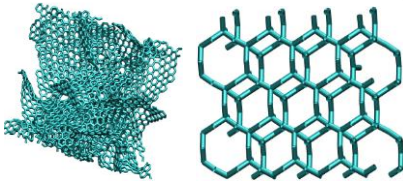
EoCoE Objectives

*Accelerate transition to **clean electricity***

Meteorology for energy : Very short term forecasting for solar and wind energy

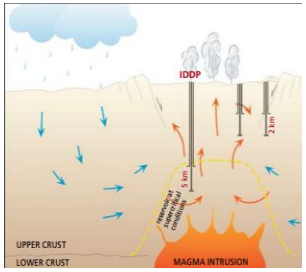
Fusion for energy : Coupling kinetic and fluid codes, mesh aligned with equilibrium configurations.



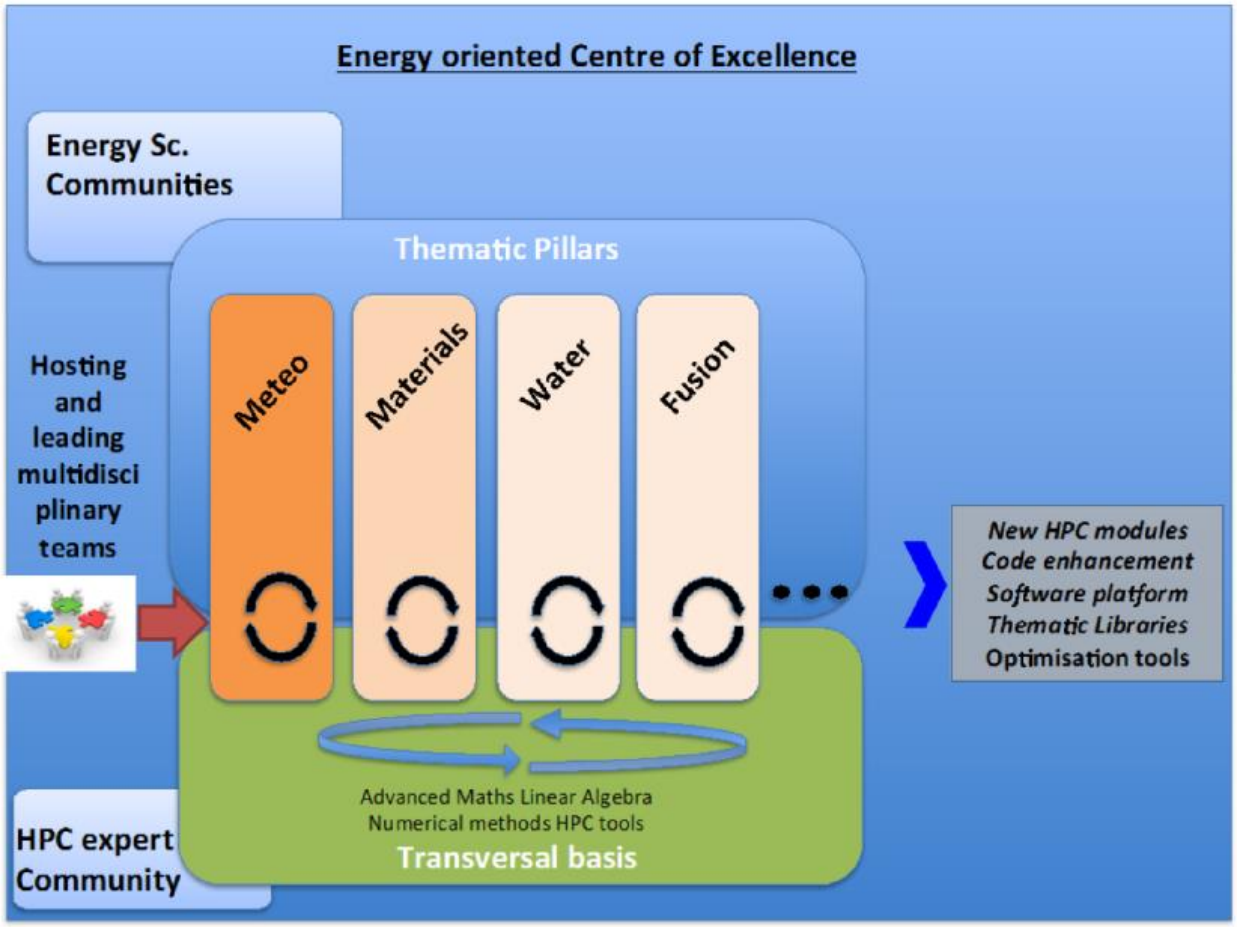


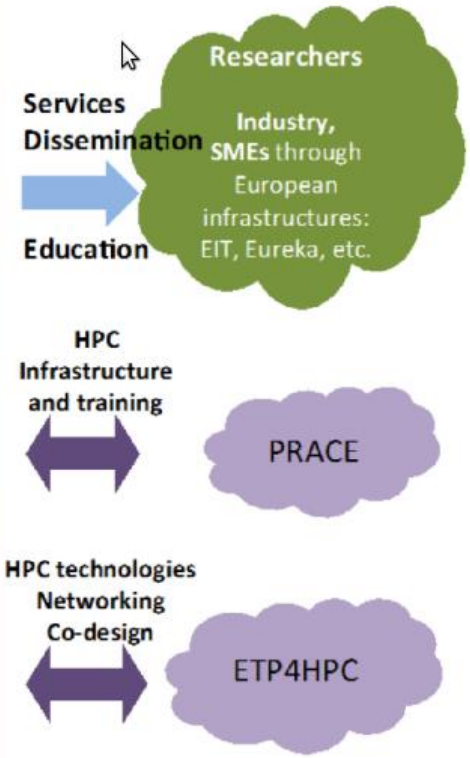
Material for energy : Photovoltaic cells, batteries and super-capacitors

Water for energy: Geothermal and hydro-power – management of resources, strategy of usage - influence of climate change.



Project structure





Consortium

Budget: 5.4 M€/3yr
Participants: 8 countries, 22 partners (9 x 3rd party)
Coordinator: Maison de la Simulation

http://www.eocoe.eu/about-us

Transversal Basis: the 5th element

EoCoE co-design objectives:

1. Tools and services for HPC (FZJ)
2. Advanced programming methods and tools (FZJ)
3. System tools for high performance (CEA)
4. Linear algebra (INRIA)
5. Numerical methods – applied mathematics (INRIA)

Integrated code optimization

- Benchmarking workshops (with POP CoE): JUBE + performance tools
- Code teams: application developers + HPC mentor
- Continual workflow: automated perf. metrics -> code perf. monitoring
- Re-engineering: tasks passed to math library/algorithm/accelerator experts

Code progress charts

Code	WP	JSC Account	Data server account	Gitlab account	JUBE integration	Benchmarks defined in JUBE	Tools integrated in JUBE	Allinea report	Score-P profile	Score-P trace	Scalasca analysis	Vampir analysis	Extrae measurement	Paraver analysis	Darshan results	VTune analysis	Advisor analysis	Performance report	Total Progress (%)	
ALYA	WP 2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	0	90	
ESIAS	WP 2	2	2	2	2	1	1	0	2	1	1	0	0	0	1	0	0	2	100	
Metalwalls	WP 3	2	2	2	2	2	2	2	2	2	2	0	2	2	2	0	2	2	100	1st workshop
PVnegf	WP 3	2	2	2	2	2	2	2	2	1	2	1	1	0	0	0	0	0	90	Dec. 2015
SHEMAT	WP 4	2	2	2	2	2	1	2	1	1	1	0	2	2	2	0	0	0	90	
ParFlow	WP 4	2	2	2	2	0	1	1	1	1	1	1	1	1	1	1	0	0	82	
GYSELA	WP 5	2	2	2	2	2	2	1	1	1	1	0	2	2	2	0	0	0	90	
nowcast system	WP 2	1	2	2	2	2	1	2	2	2	2	0	0	0	2	0	0	0	90	
CP2K	WP 3	2	2	2	2	2	0	0	1	0	0	0	1	0	2	0	0	0	64	
MDFT	WP 3	1	2	2	2	2	1	2	2	2	2	0	2	2	0	0	0	0	90	2nd workshop
TELEMAC	WP 4	1	2	2	2	2	1	2	2	2	2	2	2	2	0	0	0	0	90	May 2016
COMPASS	ext	2	2	2	2	2	1	2	2	2	2	2	2	2	0	0	0	0	90	
EIRENE	WP 5	2	2	2	2	2	1	2	2	1	1	0	1	0	2	0	0	0	90	

Legend	
0	not started
1	in progress
2	established

Code benchmarking diaries

ESIAS - EoCoE Code Diary

Code developers	Jonas Berndt (Hendrik Elbern, Charlotte Hoppe)
WP1 HPC experts	Sebastian Lührs
POP HPC experts	Brian Wylie
Location of code on EoCoE site	https://gitlab.eocoe.eu/EOCOE/Esias

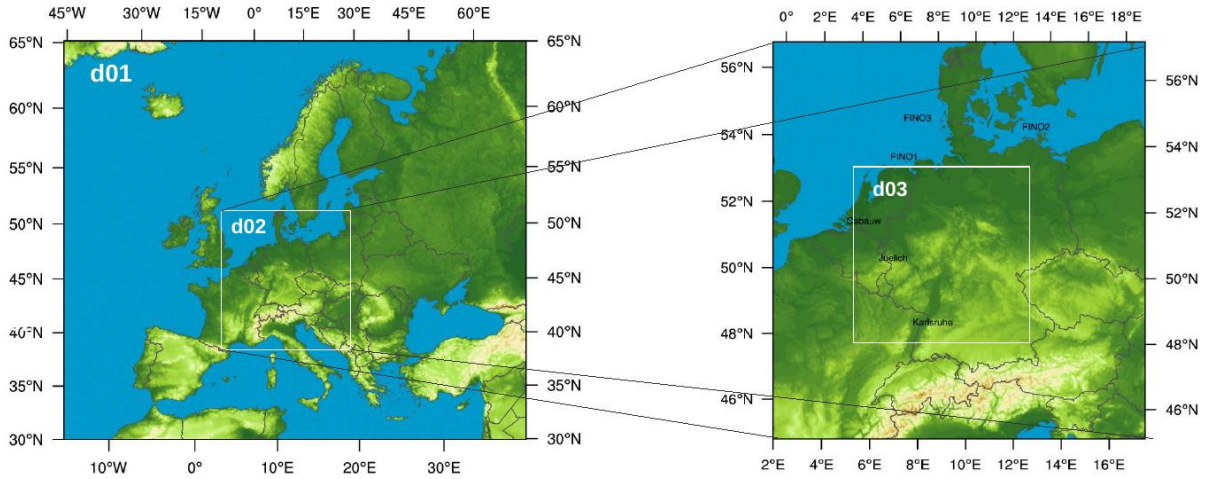
Workflow element	end of workshop 1	29/02/2016	31/03/2016	30/06/2016	30/09/2016	31/12/2016
JUBE integration	1	2	2	2		
Benchmark definition in JUBE	1	1	1	2		
Tools integrated in JUBE	0	0	1	2		
Allinea report	0	0	0	0		
Score-P profile	2	2	2	2		
Score-P trace	1					
Scalasca analysis	1					
Vampir analysis	0					
Extrae measurement	0					
Paraver analysis	0					
Darshan results	0					
VTune analysis	0					
Advisor analysis	0					

Legend	
0	not started
1	in progress
2	established

Platform				JUQUEEN			
Metric number	**Global level**	**Unit prop. MH**	**Tool(s) used**	**end of workshop 1**	**2/29/2016**	**3/31/2016**	**6/30/2016**
Glob.1	Total time elapsed	[s]	time	1,550	838	838	672
Glob.2	Total time IO	[s]	Darshan			79	87
Glob.3	Total time MPI	[s]	Scalasca	398	398	398	355
Glob.4	Memory vs compute bound	1 <= ratio <= 2					
	IO						
IO.1	IO Volume	[B]	Darshan			10,892,986,290	10,893,026,918
IO.2	Calls	[nb calls]	Darshan			110,742	96,150
IO.3	Throughput	[B] / [s]	Darshan			137,885,902	124,885,402
IO.4	Individual IO access	[B]	Darshan			309,415	181,248
	MPI						
MPI.1	P2P calls (nb)	[nb calls]					
MPI.2	P2P calls (time)	[s]	Scalasca	323	323	323	103
MPI.3	Collective calls (nb)	[nb calls]					
MPI.4	Collective calls (time)	[s]					
MPI.5	Synchro / Wait	[s]					

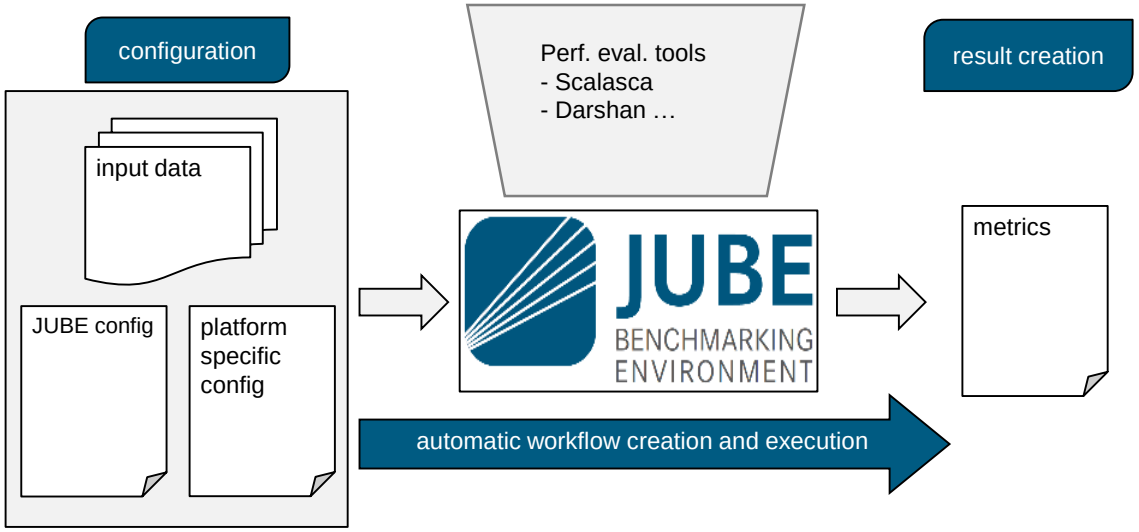
ESIAS Ultra large meteorological ensemble

- Ultra-large regional ensemble **ESIAS** with community model WRF (Weather Research & Forecasting Model), target: *O(*1000) members
- Demonstrator for detecting rare extreme error events
- Particle Filtering as an ensemble data assimilation method
- Scientific big data analytics to determine optimal perturbations



Automated metric extraction

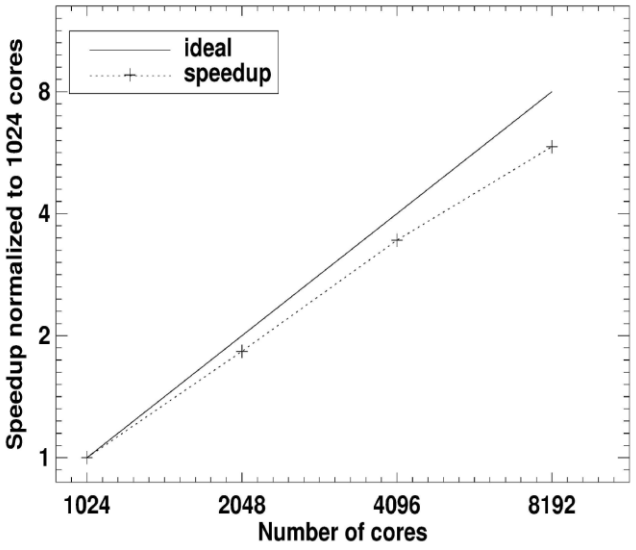
Integration of ESIAS into the EoCoE automatic metric extraction process to monitor program improvements and optimize code exchangeability

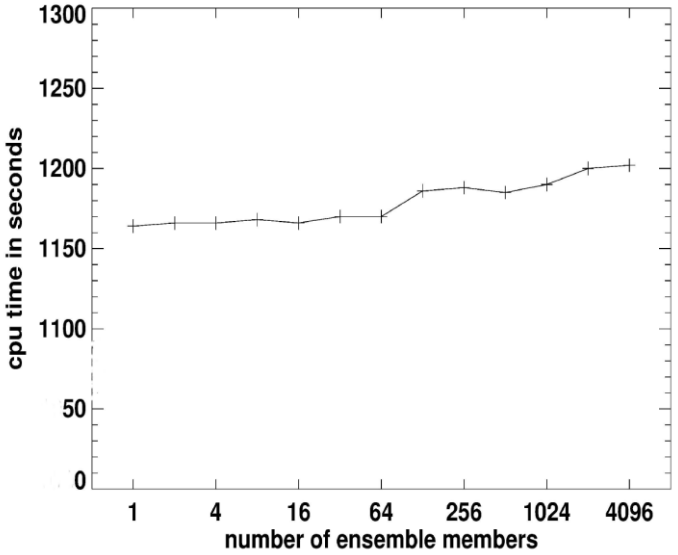


- Adaption of ESIAS to fit the JUQUEEN BlueGene Q architecture
- Evaluation of ensemble strategy
- Improvement of model runtime via compiler settings
- Evaluation of alternative I/O strategy
- Severe efficiency issues at initialization stage resolved

ESIAS code optimization in context of EoCoE

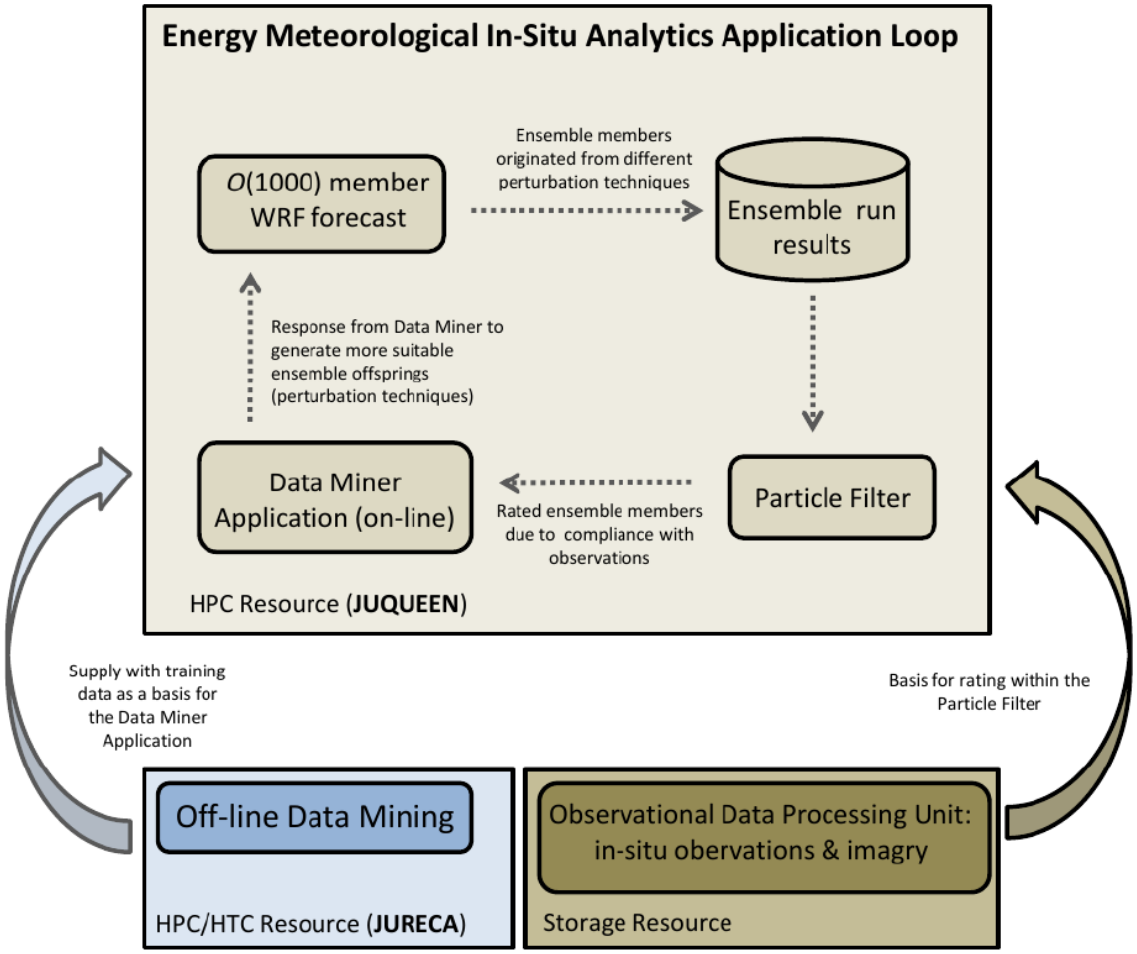
Losses through "parallel ensemble approach" are now negligible





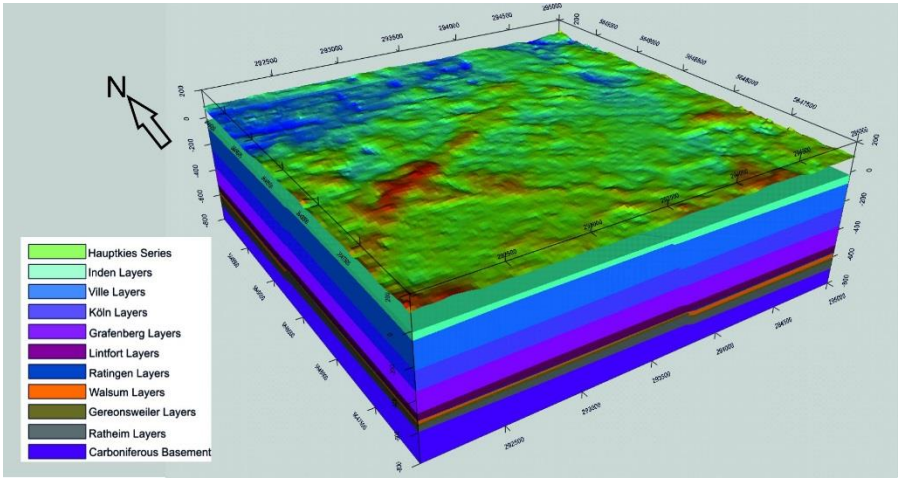
- Additional ongoing EoCoE triggered ESIAS improvements:
  Scientific big data analytics

Scientific big data analytics



- SVM (Support Vector Machine) has been ported to target application
- Off-line data mining: training of a model with historical data of different atmospheric conditions
- Initial results reach a yet unsatisfactory accuracy of 57%
- ToDo: Cross-validation (optimizing algorithm parameter) and feature engineering

SHEMAT-Suite for heat and fluid flow in porous media



Problem:

- Slow input file handling

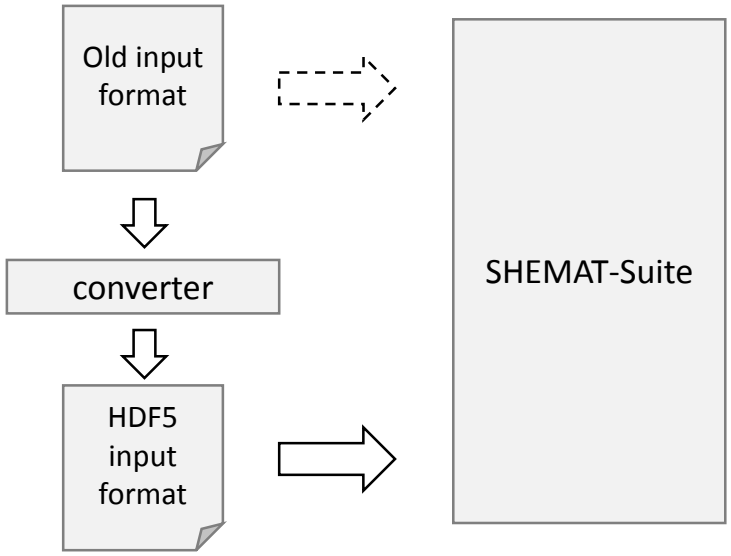
Idea:

- Switch to a different file format (HDF5)
- But: No experience in HDF5 usage and I/O optimization

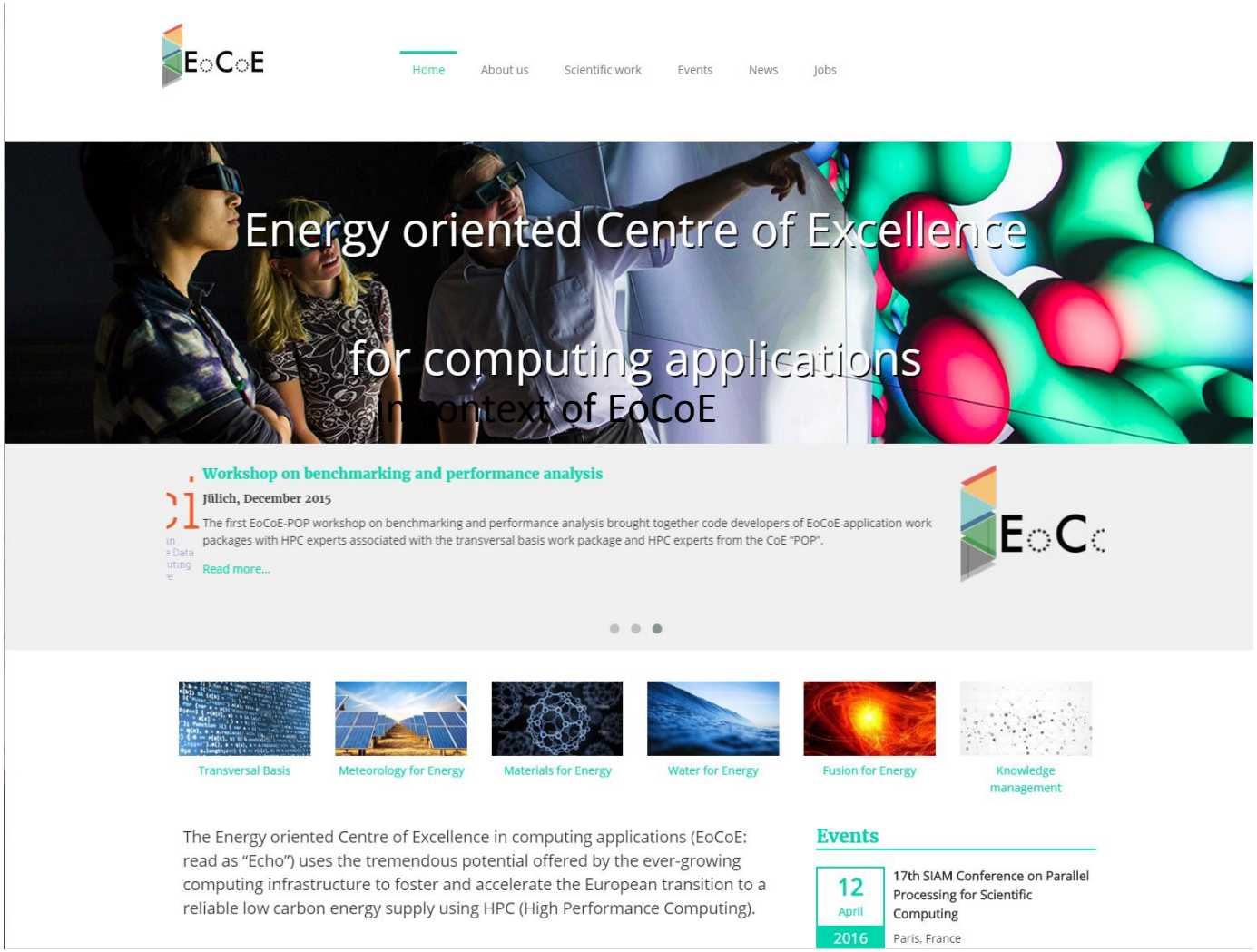
⇨ Request to EoCoE I/O experts

SHEMAT-Suite code optimization in context of EoCoE

- New SHEMAT-Suite input file handling
- Fast converter tool allows reusage of existing predecessor steps
- SHEMAT-Suite now supports new and old input format and mixed versions (for new or non converted configuration options)
- Slow input sets (> 40min input time) could be reduced to (2min input time incl. conversion process)



http://www.eocoe.eu



Thank you!

Any questions?